This is a postprint version of the following published document:

De Terán, F., Hernando, C. A note on generalized companion pencils in the monomial basis. RACSAM 114, 8 (2020).

https://doi.org/10.1007/s13398-019-00760-y

Noname manuscript No.
(will be inserted by the editor)

A note on generalized companion pencils in the monomial basis

Fernando De Terán · Carla Hernando

Received: date / Accepted: date

Abstract In this paper, we introduce a new notion of generalized companion pencils for scalar polynomials over an arbitrary field expressed in the monomial basis. Our definition is quite general and extends the notions of companion pencil in [F. De Terán, F. M. Dopico, D. S. Mackey. Linear Algebra Appl. 459 (2014) 264-333], generalized companion matrix in [C. Garnett, B. L. Shader, C. L. Shader, P. van den Driessche. Linear Algebra Appl. 498 (2016) 360-365], and Ma-Zhan companion matrices in [C. Ma, X. Zhan. Linear Algebra Appl. 438 (2013) 621-625], as well as the class of quasi-sparse companion pencils introduced in [F. De Terán, C. Hernando. INdAM Series 157-179, Springer 2019]. We analyze some algebraic properties of generalized companion pencils. We determine their Smith canonical form and we prove that they are all nonderogatory. In the last part of the work we will pay attention to the sparsity of these constructions. In particular, by imposing some natural conditions on its entries, we determine the smallest number of nonzero entries of a generalized companion pencil.

Keywords companion matrix · companion pencil · linearization · sparsity · scalar polynomial · matrix polynomial · arbitrary field · digraph · composite cycle · extension field · ring of polynomials · field of fractions.

Mathematics Subject Classification (2010) 15A22 · 65F15 · 05C20 · 05C50 · 15B99

1 Introduction

Companion matrices have been widely used to compute the roots of a monic scalar polynomial $\widetilde{q}(\lambda) = \lambda^k + \sum_{i=0}^{k-1} \lambda^i a_i$ (see, for instance, [2,3] and the references therein). Companion matrices are symbolic constructions of matrices that depend on the coefficients of the polynomial $\widetilde{q}(\lambda)$, namely $a_0, \dots, a_{k-1}$, and whose characteristic polynomial is equal to $\widetilde{q}(\lambda)$. In this way, the eigenvalues of the companion matrix are the roots of $\widetilde{q}(\lambda)$, so numerical spectral methods on the companion matrix can be used to compute the roots of $\widetilde{q}(\lambda)$.

This work has been partially supported by the *Ministerio de Economía y Competitividad* of Spain through grants MTM2017-90682-REDT and MTM2015-65798-P.

An earlier version of this paper was presented at the Conference "Linear Algebra, Matrix Analysis and Applications. ALAMA2018", held in Sant Joan d'Alacant on May/June 2018.

Fernando De Terán
Departamento de Matemáticas
Universidad Carlos III de Madrid
Avda. Universidad 30, 28911, Leganés, Spain
E-mail: fteran@math.uc3m.es

Carla Hernando
Departamento de Matemáticas
Universidad Carlos III de Madrid
Avda. Universidad 30, 28911, Leganés, Spain
E-mail: cahernan@math.uc3m.es

After many years of research in this context, a fast and backward stable method for computing the roots of monic polynomials has been recently obtained in [2]. Any scalar polynomial $q(\lambda) = \sum_{i=0}^{k} \lambda^i a_i$ (not necessarily monic) can be taken to a monic polynomial $\widetilde{q}(\lambda)$ having exactly the same roots as $q(\lambda)$ just dividing $q(\lambda)$ by its leading coefficient, a_k (namely, $\widetilde{q}(\lambda) = (1/a_k)q(\lambda)$). However, this operation may affect the accuracy of the numerical method used to compute the roots of the polynomial. For instance, the classical method for computing the spectral information of matrix polynomials [39] guarantees backward stability when the polynomial is scaled in such a way that its norm is equal to 1 (see [39, p. 576]). In order to get such a polynomial with norm equal to 1, we could divide our starting polynomial by its norm. However, this can turn a monic polynomial into a non-monic one, so this strategy is not useful if we use a numerical method which is only suitable for monic polynomials. Hence, from the numerical point of view, it may be desirable to work with non-monic polynomials (see also [3, §1]). This motivates the introduction of *companion pencils* for general (not necessarily monic) polynomials of degree k expressed in the monomial basis,

$$q(\lambda) = \sum_{i=0}^{k} \lambda^i a_i, \qquad a_i \in \mathbb{F}, \qquad i = 0, 1, \ldots, k, \qquad a_k \neq 0, \tag{1}$$

(with $\mathbb{F}$ being an arbitrary field). These companion pencils are $k \times k$ matrix pencils, $\lambda X + Y$, whose entries depend on the coefficients of the polynomial, $a_0, \ldots, a_k$, and whose determinant is equal to $q(\lambda)$ (up to a nonzero constant factor), so the roots of $q(\lambda)$ can be computed as the eigenvalues of the companion pencil. In the context of polynomial root-finding, the more recent reference [3] presents an adapted version of the algorithm in [2] for general polynomials (not necessarily monic) that uses a companion pencil (and this algorithm has been extended in [4] to matrix polynomials). The reader can also find in that reference a backward error analysis of the algorithm, together with a comparison with the algorithm that works on the companion matrix for monic polynomials.

We want to warn the reader that, despite one of the main motivations to deal with companion pencils comes from their connection to the polynomial root-finding problem, as explained above, the present work is of theoretical nature and does not consider the polynomial root-finding.

Besides the classical Frobenius companion matrices (see (2)), there is a large body of knowledge on companion matrices, mainly within the framework of the polynomial root-finding. In particular, some companion matrices for polynomials expressed in other bases than the monomial one have been introduced and used over the years [5,6,8,18,33,36,37,41]. In order to distinguish them from the classical companion matrices, some other names like "comrade", "colleague", "confederate", or "non-standard companion" have been used. The name "generalized companion matrix" (or "pencil") has been also used in many other references, with a very different meaning than the one we use in the present work, like, for instance, in [7,16,17,29,41]. However, in all these cases this notion refers to specific constructions that generalize the Frobenius companion matrix (or pencil).

Companion matrices are particular cases of companion pencils $\lambda X + Y$, where $X = I$ (and $Y = -A$, with A being the companion matrix). Companion pencils, however, allow for much more flexibility in the leading term, and, as a consequence, provide richer constructions. For instance, a companion pencil for polynomials (1) with $k = 5$ within the family of generalized Fiedler pencils [1,13] is

$$\begin{bmatrix} 0 & 0 & 0 & \lambda & \lambda a_1 + a_0 \\ 0 & 0 & 1 & 0 & -\lambda \\ 0 & \lambda & \lambda a_3 + a_2 & -1 & 0 \\ 1 & 0 & -\lambda & 0 & 0 \\ \lambda a_5 + a_4 & -1 & 0 & 0 & 0 \end{bmatrix},$$

a construction which can not be achieved with the notion of companion matrix. Note that not only the entries equal to λ are placed in non-diagonal positions, but also that some of the coefficients, besides the leading coefficient a_5, appear multiplied by λ (that is, they belong to the leading term of the pencil, which we have denoted by X).

In the monomial basis, Fiedler introduced, less than 20 years ago [28], a family of companion matrices that generalizes the Frobenius companion matrices. This family was extended to pencils, and to regular matrix polynomials, in [1], and was named in [20]. Other families of Fiedler-like pencils have been introduced in recent years, including, but not limited to, the ones in [10,11,12,13,14,21,40]. All these families are valid not just for scalar, but for matrix polynomials. The recent family of *block-Kronecker* companion pencils [25] includes all these families. Despite the efforts devoted to introduce new families of companion pencils, a comprehensive study of the structural properties of general companion pencils is still missing. Some progress in this direction has been carried out in [26, 27, 30, 35] for companion matrices. These references have analyzed interesting structural features of companion matrices (for scalar polynomials), using tools mainly from graph theory (by considering the digraph associated to the companion matrix), as well as some results and techniques from field theory.

If one looks for the most general notion of companion pencil, then it is natural to consider pencils $\lambda X + Y$ whose entries are certain functions of $a_0, \ldots, a_k$, such that $\det(\lambda X + Y) = \alpha q(\lambda)$, with $q(\lambda)$ as in (1) and $0 \neq \alpha \in \mathbb{F}$, without any further requirement (this desideratum is, for instance, rephrased in [31] for companion matrices of polynomials expressed in any basis). The set of functions of $a_0, \ldots, a_k$ containing the entries of the pencil can be as wide as desired but, in order to make the constructions practical and valid for all polynomials (1), it is natural to restrict oneself to the set of polynomials in $a_0, \ldots, a_k$. This is, precisely, the notion of generalized companion pencil we introduce in Definition 22, which is an extension, to matrix pencils, of the notion of generalized companion matrix in [30]. We use the term "generalized" because it is the one used in [30], and also in order to be consistent with previous work of the authors [19, Def. 2.2], where *companion pencil* is used for a more restrictive notion. In [35], a slightly more general notion of companion matrices is introduced, allowing the entries to be rational functions in the coefficients of the polynomial $\widetilde{q}(\lambda)$, namely $a_0, \ldots, a_{k-1}$. However, that notion includes constructions which are not valid for all coefficients $a_0, \ldots, a_{k-1}$ (namely, when the denominators of some of the rational functions vanish) except precisely in the case where all entries are just polynomials in these coefficients. A related notion of companion matrix has been also introduced in [15], but that notion is more restrictive than ours.

In [19, 26, 35], special attention has been paid to the notion of *sparsity*, namely the property of having the smallest possible number of nonzero entries. In particular, it has been proved in [35] that the smallest number of nonzero entries in a companion matrix for scalar polynomials of degree k is $2k-1$, and, in [26], a canonical expression (up to permutation similarity) has been presented for the so-called sparse *companion patterns*. Up to our knowledge, the only reference containing a study on similar issues for companion pencils is the recent work [19]. In that reference, a canonical expression (up to permutation equivalence) is also presented for a general class of so-called *quasi-sparse* companion pencils, which resembles the class of companion patterns in [26]. Using this canonical expression, the number of different sparse companion pencils in that class has been also determined.

The present work is a complement to [19]. In particular, we deal not only with companion pencils (including those in the family introduced in [19]), but also with *generalized companion pencils* (see Definitions 21 and 22), where we allow the non-constant entries of the pencil to be not just the coefficients, a_i, of the polynomial (1), but polynomials in these coefficients. We are also interested in the sparsity of these constructions, namely in determining the smallest possible number of nonzero entries (these issues are addressed in Section 5). This may seem to be, a priori, a problem easy to solve once the solution for companion matrices is known. However, in the case of companion matrices, $\lambda I - A$, the leading term (the identity) is fixed, so one can disregard this term and just look for nonzero entries in the trailing coefficient A. By contrast, for companion pencils, $\lambda X + Y$, none of X and Y are fixed, so the nonzero terms of $\lambda X + Y$ come from nonzero terms in either X, or Y, or both. This is the main reason why the arguments for the case of companion matrices cannot be extended to generalized companion pencils in a straightforward way.

In Section 4 we impose the condition that each coefficient a_i appears only once in the generalized companion pencil, and we determine the structure of the nonzero entries of the pencil in this case. This condition will be imposed again in Section 5 to obtain sparse generalized companion pencils (see Theorem 53). To impose this condition is very natural, since, if we look for the generalized companion pencil to include the smallest amount of information, it is natural to avoid repetitions of the coefficients.

In Section 3, we characterize the set of generalized companion pencils by using the Smith canonical form. More precisely, we show that all generalized companion pencils of the polynomial (1) have the same Smith form over the field of rational functions in $a_0, \ldots, a_k$, namely $\operatorname{diag}(I_{k-1}, \frac{1}{a_k}q(\lambda))$ (Theorem 31). We also prove that all generalized companion pencils are *nonderogatory* (Theorem 32).

2 Basic definitions

Let $\mathbb{F}$ be an arbitrary field and let $z_0, \ldots, z_k$ be distinct indeterminates. We denote by $\mathbb{F}[z_0, \ldots, z_k]$ the ring of polynomials in $z_0, \ldots, z_k$ over $\mathbb{F}$, and by $\mathbb{F}(z_0, \ldots, z_k)$ the field of rational functions of $z_0, \ldots, z_k$ over $\mathbb{F}$, namely

$$\mathbb{F}(z_0, \ldots, z_k) := \left\{ \frac{f}{g} : \ f, g \in \mathbb{F}[z_0, \ldots, z_k], \ g \neq 0 \right\}.$$

The notion of companion pencil for matrix polynomials recently introduced in [19, Def. 2.2] (see also [23, p. 296]) imposes a severe restriction on the nonzero blocks to be either the identity matrix or one of the coefficients of the polynomial (multiplied by some constant in $\mathbb{F}$). This restriction was motivated for practical purposes, and also because most of the families of companion pencils introduced at that time in the literature (including the Fiedler-like families) enjoyed that property. Even though we are not going to work with this notion, we include here, for the sake of completeness, its particularization for scalar polynomials.

Definition 21 *A* companion pencil *for polynomials* (1) *of degree k is a $k \times k$ matrix pencil $L(\lambda) = \lambda X + Y$ such that:*

(i) *each nonzero entry of X and Y is either a scalar (i.e., an element in $\mathbb{F}$) or a scalar multiplied by a_i, for some $i = 0, \ldots, k$, and*
(ii) $\det L(\lambda) = \alpha q(\lambda)$*, for some $0 \neq \alpha \in \mathbb{F}$.*

However, when looking for the most general notion of companion pencil, it is natural to consider pencils as in Definition 22. It should be noted that this notion is motivated by the definition of *Ma-Zhan companion matrices* in [35] (see [30], where this name was introduced), which are companion matrices for monic polynomials whose entries belong to $\mathbb{F}(a_0, \ldots, a_{k-1})$. In this work, however, we are interested in practical constructions, valid for all polynomials, so we restrict ourselves to entries in the ring $\mathbb{F}[a_0, \ldots, a_k]$.

Definition 22 *A* generalized companion pencil *for polynomials* (1) *of degree k is a $k \times k$ matrix pencil $L(\lambda) = \lambda X + Y$ such that:*

(i) *each nonzero entry of X and Y belongs to $\mathbb{F}[a_0, \ldots, a_k]$, and*
(ii) $\det L(\lambda) = \alpha q(\lambda)$*, for some $0 \neq \alpha \in \mathbb{F}$.*

Definition 22 arises when looking for more general constructions than the ones in Definition 21, allowing for the nonzero entries to be in the most general set that includes the coefficients of the polynomial and is valid for all polynomials (1), namely the ring of polynomials in $a_0, \ldots, a_k$. In particular, Definition 22 extends Definition 21, together with the class of generalized companion matrices in [30], and the class of Ma-Zhan matrices in [35]. The notions of generalized companion matrix in [30] and of Ma-Zhan matrix in [35] are only introduced for monic scalar polynomials. In the generalized companion matrices in [30], the nonzero entries are placed in some specific part of the matrix (namely, they are on or below the superdiagonal), and some of them are prescribed (namely, all entries in the superdiagonal are equal to 1). Also, in the class of Ma-Zhan matrices, the number of nonzero entries is restricted to be $2k - 1$ (note, however, that if we look at companion matrices as particular cases of companion pencils, we need to add another k nonzero entries coming from λI_k, see below). By contrast, our definition of generalized companion pencil does not specify any single entry, nor imposes any condition on the number of nonzero entries. However, in Section 5, we are interested in *sparse* generalized companion pencils, namely, those having the smallest number of nonzero entries.

The following example shows one companion pencil and one generalized companion pencil which is not companion, according to (i) in Definition 21.

Example 21 *Let $q(\lambda) = \sum_{i=0}^{3} \lambda^i a_i$ be a polynomial of degree* 3*. Let us consider the following matrix pencils $L_1(\lambda)$ and $L_2(\lambda)$:*

$$L_1(\lambda) = \begin{bmatrix} \lambda a_3 + a_2 & -1 & 0 \\ 0 & \lambda & -1 \\ \lambda a_1 + a_0 & 0 & \lambda \end{bmatrix} \quad \text{and} \quad L_2(\lambda) = \begin{bmatrix} \lambda a_1 a_2 + a_0 & -a_1 a_2 & \lambda \\ \lambda & -1 & 0 \\ \lambda a_2 + a_1 & \lambda a_3 & -1 \end{bmatrix}.$$

In both cases, $\det L_1(\lambda) = \det L_2(\lambda) = \alpha \sum_{i=0}^{3} \lambda^i a_i$*, with $\alpha = 1$ (condition* (ii) *in* Definitions 21 and 22*), so $L_1(\lambda)$ is a companion pencil and $L_2(\lambda)$ is a generalized companion pencil for $q(\lambda)$.*

We note that condition (ii) in both Definitions 21 and 22 is the natural extension of the identity $\det(\lambda I_k - A) = \widetilde{q}(\lambda)$ for a companion matrix A, with $\widetilde{q}(\lambda) = \lambda^k + \sum_{i=0}^{k-1} \lambda^i a_i$. Therefore, companion matrices are particular cases of (generalized) companion pencils (namely, when $L(\lambda) = \lambda I_k - A$).

3 Some properties of generalized companion pencils

Matrix pencils (and, in particular, companion pencils) are particular cases of *matrix polynomials*, that is, matrices with polynomial entries. One of the standard transformations acting on matrix polynomials is the *unimodular equivalence* transformation. More precisely, given a field $\mathbb{K}$, two matrices $A(\lambda)$ and $B(\lambda)$ with entries in $\mathbb{K}[\lambda]$ are said to be *unimodularly equivalent* over $\mathbb{K}$ if there exist two unimodular matrices $U(\lambda)$ and $V(\lambda)$ such that $B(\lambda) = U(\lambda)A(\lambda)V(\lambda)$. We recall that $M(\lambda) \in \mathbb{K}[\lambda]^{m\times m}$ is *unimodular* if $\det M(\lambda)$ is a nonzero constant in $\mathbb{K}$. The canonical form for matrix polynomials under unimodular transformations is the *Smith form*, that we recall here for the sake of completeness (see, for instance, [32, Chapter S1]).

Definition 31 (Smith form). *Let $Q(\lambda)$ be an $m \times m$ matrix polynomial over an arbitrary field $\mathbb{K}$. Then there exist two $m \times m$ unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ over $\mathbb{K}$ such that*

$$U(\lambda)Q(\lambda)V(\lambda) = \operatorname{diag}(d_1(\lambda), \dots, d_m(\lambda)) =: D(\lambda),$$

where $d_i(\lambda) \in \mathbb{K}[\lambda]$, for $i = 1, \dots, m$, are monic polynomials and they form a divisibility chain, that is, $d_j(\lambda)$ is a divisor of $d_{j+1}(\lambda)$ (denoted by $d_j(\lambda)|d_{j+1}(\lambda)$), for $j = 1, \dots, m-1$. The $m\times m$ diagonal matrix $D(\lambda)$ is unique and is called the Smith form of $Q(\lambda)$ over $\mathbb{K}$.

For generalized companion pencils over a field $\mathbb{F}$, the field $\mathbb{K}$ above is the field of rational functions $\mathbb{F}(a_0, \dots, a_k)$. All (generalized) companion pencils $\lambda X + Y$ introduced in [1,12,13,25] are *linearizations* of $q(\lambda)$, which means that they are unimodularly equivalent to $\operatorname{diag}(I_{k-1}, q(\lambda))$. Moreover, the unimodular transformations leading to $\operatorname{diag}(I_{k-1}, q(\lambda))$ can be chosen in such a way that their determinant belongs to the base field $\mathbb{F}$ (i.e., it is independent of the coefficients $a_0, \dots, a_k$). In other words, the Smith canonical form over $\mathbb{F}$ of all these pencils is $\operatorname{diag}(I_{k-1}, \frac{1}{a_k} q(\lambda))$ (notice that, in the Smith canonical form, all nonzero entries are monic polynomials). The key for doing this is the presence of $k-1$ entries equal to 1 (though some nonzero entries in the field $\mathbb{F}$ would be enough), placed in some specific positions which allow one to find some unimodular transformations that take $\lambda X + Y$ into $\operatorname{diag}(I_{k-1}, q(\lambda))$. As a consequence, for each particular value of the coefficients $a_0, \dots, a_k$, all these companion pencils are unimodularly equivalent over $\mathbb{F}$ to both the *first* and *second Frobenius companion pencils*, $F_1(\lambda) = \lambda \operatorname{diag}(a_k, I_{k-1}) - C_1$ and $F_2(\lambda) = \lambda \operatorname{diag}(a_k, I_{k-1}) - C_2$, with

$$C_1 = \begin{bmatrix} -a_{k-1} & -a_{k-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad C_2 = C_1^T. \tag{2}$$

In other words, these companion pencils can be taken to $F_1(\lambda)$ and $F_2(\lambda)$ by means of unimodular transformations over $\mathbb{F}(a_0, \dots, a_k)$ whose determinant does not depend on $a_0, \dots, a_k$. We are going to see that every generalized companion pencil is also unimodularly equivalent over $\mathbb{F}(a_0, \dots, a_k)$ to $\operatorname{diag}(I_{k-1}, q(\lambda))$. This is a strong result, since the only condition we impose on generalized companion pencils is that their determinant is equal to $q(\lambda)$ (up to a nonzero constant factor). However, the proof, as we will see, is quite elementary, and it is based on the following lemma, which is a straightforward generalization of Lemma 2 in [35] from monic polynomials to general (not necessarily monic) polynomials.

Lemma 31 *Let $\mathbb{F}$ be a field and let $a_0, \dots, a_k$ be distinct indeterminates. Then, the polynomial $q(a_0, \dots, a_k; \lambda) = \sum_{i=0}^{k} \lambda^i a_i$ is irreducible over $\mathbb{F}(a_0, \dots, a_k)$.*

Proof For brevity, we write $q(\lambda) := q(a_0, \ldots, a_k; \lambda)$. Let us assume, by contradiction, that

$$q(\lambda) = q_1(\lambda) q_2(\lambda), \tag{3}$$

with $\deg(q_1)$, $\deg(q_2) \geq 1$.

If, for a given polynomial $q(a_0, \ldots, a_k; \lambda)$, we define the monic polynomial $\widetilde{q}(\lambda) := q(a_0, \ldots, a_{k-1}, 1; \lambda)$, then:

$$\widetilde{q}(\lambda) = \widetilde{q_1}(\lambda) \widetilde{q_2}(\lambda), \tag{4}$$

where $\widetilde{q_1}(\lambda), \widetilde{q_2}(\lambda) \in \mathbb{F}(a_0, \ldots, a_{k-1})$.

Note that Eq. (4) is obtained from (3) after replacing $a_k = 1$ in $q(\lambda)$, $q_1(\lambda)$, and $q_2(\lambda)$. Now, we prove that $\widetilde{q}(\lambda)$ is reducible. For this, it suffices to prove that both $\widetilde{q_1}(\lambda)$ and $\widetilde{q_2}(\lambda)$ have degree, at least, 1 in λ. Since $\deg(\widetilde{q_1}) + \deg(\widetilde{q_2}) = k$, if $\deg(\widetilde{q_1}) = 0$ (respectively, $\deg(\widetilde{q_2}) = 0$) it would be $\deg(\widetilde{q_2}) = k$ (respectively, $\deg(\widetilde{q_1}) = k$), but this is a contradiction with the fact that

$$\deg(\widetilde{q_1}) \leq \deg(q_1) \leq k-1, \quad \text{and} \quad \deg(\widetilde{q_2}) \leq \deg(q_2) \leq k-1.$$

Therefore, (4) implies that $\widetilde{q}(\lambda) = \lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_1\lambda + a_0$ is reducible over $\mathbb{F}(a_0, \ldots, a_{k-1})$, in contradiction with Lemma 2 in [35]. □

Theorem 31 *Let $q(\lambda)$ be a polynomial* (1) *of degree k with $a_0, \ldots, a_k$ being distinct indeterminates. Let $L(\lambda) = \lambda X + Y$ be a generalized companion matrix pencil of $q(\lambda)$. Then, the Smith form of $L(\lambda)$ over $\mathbb{F}(a_0, \ldots, a_k)$ is* $\operatorname{diag}(I_{k-1}, \frac{1}{a_k} q(\lambda))$.

Proof Let $L(\lambda)$ be a generalized companion matrix pencil of $q(\lambda)$ (so $\det L(\lambda) = \alpha q(\lambda)$, for some $0 \neq \alpha \in \mathbb{F}$), and let $D(\lambda)$ be the Smith form of $L(\lambda)$ over $\mathbb{F}(a_0, \ldots, a_k)$ (see Definition 31). Then,

$$U(\lambda) L(\lambda) V(\lambda) = \begin{bmatrix} d_1(\lambda) & & \\ & \ddots & \\ & & d_k(\lambda) \end{bmatrix} =: D(\lambda),$$

where $d_i(\lambda)$, for $i = 1, \ldots, k$, are monic polynomials, and $d_1(\lambda) | d_2(\lambda) | \cdots | d_k(\lambda)$ over $\mathbb{F}(a_0, \ldots, a_k)$.

Taking determinants in the previous identity, we obtain that

$$\beta q(\lambda) = d_1(\lambda) \cdots d_k(\lambda),$$

with $0 \neq \beta := \alpha \det U(\lambda) \det V(\lambda) \in \mathbb{F}(a_0, \ldots, a_k)$.

By Lemma 31, $q(\lambda)$ is irreducible over $\mathbb{F}(a_0, \ldots, a_k)$. This implies, by the divisibility conditions in the polynomials d_i, together with the fact that they are all monic, that $d_1(\lambda) = \cdots = d_{k-1}(\lambda) = 1$ and $d_k(\lambda) = \frac{1}{a_k} q(\lambda)$, as wanted. □

Note, however, that the determinant of the unimodular transformations $U(\lambda)$ and $V(\lambda)$ in the proof of Theorem 31 taking $\lambda X + Y$ to its Smith form belongs to the field $\mathbb{F}(a_0, \ldots, a_k)$ so, in principle, the transformations could not be defined for some values of the coefficients $a_0, \ldots, a_k$.

Another relevant feature of all companion matrices (or pencils) for polynomials given in the monomial basis known so far is the property of being *nonderogatory* (that is, all their eigenvalues have geometric multiplicity equal to 1). This property also holds for companion matrices of polynomials expressed in other bases introduced in the literature (see, for instance, [6, p. 279] and [15, Th. 4.1]). For companion pencils (and matrices) in the monomial basis, the property is an immediate consequence of the fact that they are unimodularly equivalent over $\mathbb{F}[a_0, \ldots, a_k]$ to the first (and the second) Frobenius companion pencil (or matrix), and to the presence of the $k-1$ entries equal to 1 in the Frobenius companion matrices C_1 and C_2 (2). Here it is important to emphasize that the unimodular equivalence can be performed using matrices with entries in the ring $\mathbb{F}[a_0, \ldots, a_k]$ and whose determinant is a constant in $\mathbb{F}$, which guarantees that they are well defined and preserve the rank, regardless of the values of $a_0, \ldots, a_k$. However, for arbitrary generalized companion pencils this cannot be guaranteed since, as mentioned before, the unimodular transformations could not be defined for some values of these coefficients. However, the following result shows that this property of being nonderogatory holds for any generalized companion pencil of scalar polynomials over any infinite field.

Theorem 32 *Let $\mathbb{F}$ be an infinite field, and let $q(\lambda)$ be a polynomial* (1) *of degree k, with $a_0, \ldots, a_k$ being distinct indeterminates. Let $L(\lambda) = \lambda X + Y$ be a generalized companion matrix pencil of $q(\lambda)$. Then, for all $a_0, \ldots, a_k \in \mathbb{F}$, $L(\lambda)$ is nonderogatory.*

Proof In order to emphasize that both the polynomial $q(\lambda)$ and the generalized companion pencil $L(\lambda)$ in the statement depend on the coefficients of the polynomial, we write them as $q(a_0, \ldots, a_k; \lambda)$ and $L(a_0, \ldots, a_k; \lambda)$, respectively, throughout this proof. We are going to prove that $\operatorname{rank} L(a_0, \ldots, a_k; \lambda_0) \geq k-1$, for all $\lambda_0 \in \mathbb{F}$ and all $(a_0, \ldots, a_k) \in \mathbb{F}^{k+1}$.

If $q(a_0, \ldots, a_k; \lambda_0) = \sum_{i=0}^{k} \lambda_0^i a_i \neq 0$, then $\operatorname{rank} L(a_0, \ldots, a_k; \lambda_0) = k$, since $\det L(a_0, \ldots, a_k; \lambda_0) \neq 0$.

Now, let $\lambda_0 \in \mathbb{F}$ be such that $q(a_0, \ldots, a_k; \lambda_0) = 0$, for some $(a_0, \ldots, a_k) \in \mathbb{F}^{k+1}$. Let us write:

$$L(a_0, \ldots, a_k; \lambda_0) = q(a_0, \ldots, a_k; \lambda_0) L_0(a_0, \ldots, a_k) + L_1(a_1, \ldots, a_k), \tag{5}$$

where $L_1(a_1, \ldots, a_k)$ does not depend on a_0. Such an expression is always possible after writing

$$a_0 = q(a_0, \ldots, a_k; \lambda_0) + (a_0 - q(a_0, \ldots, a_k; \lambda_0)) = q(a_0, \ldots, a_k; \lambda_0) + h(a_0, \ldots, a_k; \lambda_0),$$

where $h(a_0, \ldots, a_k; \lambda_0) := \sum_{i=1}^{k} \lambda_0^i a_i$ does not depend on a_0. If $\operatorname{rank} L_1(a_1, \ldots, a_k) \geq k-1$, for all $(a_1, \ldots, a_k)$, then we are done, since for those $(a_0, \ldots, a_k)$ such that $q(a_0, \ldots, a_k; \lambda_0) = 0$ we have $\operatorname{rank} L(a_0, \ldots, a_k; \lambda_0) = \operatorname{rank} L_1(a_1, \ldots, a_k)$. Suppose, by contradiction, that there is some $(a_0^{(0)}, \ldots, a_k^{(0)})$ such that $q(a_0^{(0)}, \ldots, a_k^{(0)}; \lambda_0) = 0$ and $\operatorname{rank} L_1(a_1^{(0)}, \ldots, a_k^{(0)}) < k-1$. Set $\widehat{q}(a_0) := q(a_0, a_1^{(0)}, \ldots, a_k^{(0)}; \lambda_0)$, which is a polynomial in a_0 with coefficients in $\mathbb{F}$. Now, for all $a_0 \in \mathbb{F}$, we have

$$\begin{aligned} \widehat{q}(a_0) &= \alpha \det L(a_0, a_1^{(0)}, \ldots, a_k^{(0)}; \lambda_0) \\ &= \alpha \det \left(\widehat{q}(a_0) L_0(a_0, a_1^{(0)}, \ldots, a_k^{(0)}) + L_1(a_1^{(0)}, \ldots, a_k^{(0)}) \right) \\ &= \alpha\, \widehat{q}(a_0)^2 \cdot p(a_0), \end{aligned} \tag{6}$$

where $p(a_0)$ is a polynomial in a_0 with coefficients in $\mathbb{F}$ (and $0 \neq \alpha \in \mathbb{F}$). To see the last identity in (6), set $L_1 := L_1(a_1^{(0)}, \ldots, a_k^{(0)}) \in \mathbb{F}^{k \times k}$. Since $\operatorname{rank} L_1 < k-1$, all $(k-1) \times (k-1)$ minors of L_1 are zero. Now, the formula for the determinant of a sum of two matrices in [38, Th. 4.1] allows us to conclude that $\det \left(\widehat{q}(a_0) L_0(a_0, a_1^{(0)}, \ldots, a_k^{(0)}) + L_1(a_1^{(0)}, \ldots, a_k^{(0)}) \right)$ is a multiple of $\widehat{q}(a_0)^2$ (the compound matrices $\mathcal{C}_k(L_1) = \det L_1$ and $\mathcal{C}_{k-1}(L_1)$ in that formula are zero). Then, (6) implies that $\widehat{q}(a_0)(\widehat{q}(a_0)p(a_0) - 1)$, which is a non-identically zero polynomial in a_0, vanishes over $\mathbb{F}$. Since $\mathbb{F}$ is an infinite field, this is impossible. □

Theorem 32 is, again, a quite strong result, since the only condition imposed to a pencil with entries in $\mathbb{F}[a_0, \ldots, a_k]$ for being a generalized companion pencil of $q(\lambda)$ is that its determinant coincides, up to a nonzero constant factor, with the given polynomial $q(\lambda)$.

During the refereeing process of this manuscript we have been informed by K. N. Vander Meulen on the recent reference [24], that deals with companion matrices for real monic polynomials, and where it is proved ([24, Theorem 3.1]) that they are nonderogatory. However, we want to emphasize that the notion of companion matrix used in that reference imposes the restriction that each coefficient a_i appears just once in the matrix, and that the remaining entries are constant.

4 Generalized companion pencils where each coefficient appears only once

In [19], we looked for companion pencils (as in Definition 21) with a small number of nonzero entries. There, however, only nonzero entries of the form a_i, λa_{i+1} or $\lambda a_{i+1} + a_i$ (up to scalar constants) were considered, and this led to the class of quasi-sparse pencils with at most $3k-2$ nonzero entries, denoted by $\mathcal{R}_{n,k}$ (see [19, Def. 2.3]).

Proposition 41 shows that the nonzero entries mentioned in the precedent paragraph are, up to constants, the nonzero entries in a generalized companion pencil when we impose the restriction that each coefficient a_i, for $i = 0, \ldots, k$, of the polynomial (1) appears in just one entry.

Proposition 41 *Let $L(\lambda) = \lambda X + Y$ be a generalized companion pencil for polynomials* (1) *of degree k such that each coefficient a_i, for $i = 0, \ldots, k$, appears in just one entry of $L(\lambda)$. Then the entry containing a_i is either of the form*

$$\lambda(\beta_1 a_{i+1} + \beta_2) + \alpha_1 a_i + \alpha_2, \quad \text{for} \quad 0 \leq i \leq k-1, \tag{7}$$

with $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{F}$, and $\alpha_1 \neq 0$, or

$$\lambda(\beta_1 a_i + \beta_2) + \alpha_1 a_{i-1} + \alpha_2, \quad \text{for} \quad 0 \leq i \leq k-1, \tag{8}$$

with $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{F}$, and $\beta_1 \neq 0$.

Proof Since $L(\lambda)$ is a generalized companion pencil for $q(\lambda)$, its (s,t) entry is of the form $\ell_{st}(\lambda) = \lambda x_{st} + y_{st}$, for $s, t = 1, \ldots, k$, with $x_{st}, y_{st} \in \mathbb{F}[a_0, \ldots, a_k]$. By hypothesis, a_i, for $0 \leq i \leq k$, can be either in x_{st} or y_{st} for just one (s,t). Now, we claim that:

1. If a_i is in x_{st}, then: (a) a_i does not appear in y_{st}. (b) There is no other a_j in x_{st}. (c) The only possible a_j appearing in y_{st} is a_{i-1}. (d) x_{st} must be equal to $\beta_1 a_i + \beta_2$, for some $\beta_1, \beta_2 \in \mathbb{F}$, with $\beta_1 \neq 0$. (e) y_{st} is of the form $\alpha_1 a_{i-1} + \alpha_2$, for some $\alpha_1, \alpha_2 \in \mathbb{F}$.
2. Similarly, if a_i is in y_{st}, then: (a') a_i does not appear in x_{st}. (b') There is no other a_j in y_{st}. (c') The only possible a_j appearing in x_{st} is a_{i+1}. (d') y_{st} must be equal to $\alpha_1 a_i + \alpha_2$, for some $\alpha_1, \alpha_2 \in \mathbb{F}$, with $\alpha_1 \neq 0$. (e') x_{st} is of the form $\beta_1 a_{i+1} + \beta_2$, for some $\beta_1, \beta_2 \in \mathbb{F}$.

Note that case 1 above corresponds to (8), whereas case 2 corresponds to (7). Then, it remains to prove (a)-(e) and (a')–(e'). We assume that a_i is in x_{st}. Condition (*ii*) in Definition 22 implies $\det L(\lambda) = \alpha q(\lambda)$, with $0 \neq \alpha \in \mathbb{F}$, and $q(\lambda)$ as in (1). Spanning the determinant either across the sth row or the tth column, we get $\det L(\lambda) = \ell_{st} C_{st} + q_{st}(\lambda)$, where ℓ_{st} is the (s,t) entry of $L(\lambda)$, C_{st} is the cofactor of the entry containing ℓ_{st}, and $q_{st}(\lambda)$ does not contain a_i. Then, setting $\ell_{st} = \lambda x_{st} + y_{st}$, we have

$$\det L(\lambda) = (\lambda x_{st} + y_{st}) C_{st} + q_{st}(\lambda), \tag{9}$$

and C_{st} is of the form $C_{st} = \lambda^{i-1} c_{i-1} + \cdots + \lambda^{i-r} c_{i-r}$, for some $r \geq 1$, and where $c_{i-1}, \ldots, c_{i-r}$ are polynomials in the coefficients a_j, for $j \neq i$, and $c_{i-r} \neq 0$. Note that, since a_i is in x_{st}, c_{i-1} is always a nonzero polynomial as well. Note also that the degree of all possible appearances of a_i in ℓ_{st} must be equal to 1. If, on the contrary, there is a term a_i^s, with $s > 1$, in ℓ_{st}, by (9) and taking into account that a_i does not appear in either C_{st} and $q_{st}(\lambda)$, this term a_i^s would appear in $\det L(\lambda)$, a contradiction with the fact that $\det L(\lambda) = \alpha q(\lambda)$, with $0 \neq \alpha \in \mathbb{F}$.

We consider two cases, depending on whether a_i is in y_{st} or not.

– If a_i is only in x_{st}, then $C_{st} = \gamma \lambda^{i-1}$, for some $\gamma \in \mathbb{F} \setminus \{0\}$, since otherwise there would be another term containing a_i in $\det L(\lambda)$, besides $\lambda^i a_i$ (recall that a_i does not appear in either C_{st} and $q_{st}(\lambda)$). Hence, no other coefficient a_j than a_{i-1} may appear in ℓ_{st} since, otherwise, there would be a nonzero term of the form $\lambda^{i-1} a_j$ or $\lambda^i a_j$ in $\det L(\lambda)$ (recall, again, that a_j does not appear in either C_{st} and $q_{st}(\lambda)$). Moreover, following similar arguments as those for a_i, the coefficient a_{i-1} must appear in y_{st}, also with degree 1, and cannot appear in x_{st}. Thus, we have proved claims (a)–(c). For claims (d) and (e) just note that, by the previous considerations, x_{st} and y_{st} are linear polynomials in, respectively, a_i and a_{i-1} over $\mathbb{F}$. Therefore, they are as claimed in (d) and (e), respectively.
– If a_i appears in both x_{st} and y_{st}, then $\ell_{st} = \lambda \beta_1 a_i + \alpha_1 a_i + \widetilde{\ell}_{st}$, for some $\alpha_1, \beta_1 \in \mathbb{F} \setminus \{0\}$, where a_i does not appear in $\widetilde{\ell}_{st}$, and we would have, from (9),

$$\begin{aligned} \det L(\lambda) &= (\lambda \beta_1 a_i + \alpha_1 a_i + \widetilde{\ell}_{st})(\lambda^{i-1} c_{i-1} + \cdots + \lambda^{i-r} c_{i-r}) + q_{st}(\lambda) = \\ &= a_i(\lambda^i \beta_1 c_{i-1} + \lambda^{i-r} \alpha_1 c_{i-r}) + \widehat{q}_{st}(\lambda), \end{aligned} \tag{10}$$

where $\widehat{q}_{st}(\lambda)$ is another polynomial in λ with coefficients in $\mathbb{F}[a_0, \ldots, a_k]$. The relevant property of $\widehat{q}_{st}(\lambda)$ is that the terms in this polynomial containing a_i (if any) are of degree larger than $i - r$ and smaller than i in λ, so the first term in the second line of the right-hand side of (10) necessarily appears in $\det L(\lambda)$ (note that this also holds if $r = 1$). In other words, there is a term of degree $i - r$ in λ containing a_i, in contradiction with the fact that $\det L(\lambda) = \alpha q(\lambda)$, with $0 \neq \alpha \in \mathbb{F}$.

This proves (a)-(e), and claims (a')-(e') can be proved in a similar way. □

5 Sparse generalized companion pencils

In this section, we are interested in generalized companion pencils having the smallest number of nonzero entries. By a *nonzero entry* we mean an entry of the form $\lambda x + y$, where x, y are polynomials in $a_0, \ldots, a_k$ and at least one of them is non-identically zero (that is, it has, at least, one nonzero coefficient). In Theorem 51 we get a lower bound on the number of nonzero entries in the coefficients, X and Y, of a generalized companion pencil $L(\lambda) := \lambda X + Y$. There, the total number of nonzero entries is obtained after adding up the nonzero entries of X with those of Y. However, some of these nonzero entries can be placed in the same position in both X and Y, so the number of nonzero entries of $L(\lambda)$ can be smaller than this lower bound. By imposing some additional restrictions on X and Y we can get a sharper bound, in Theorems 52 and 53, for the number of nonzero entries in $L(\lambda)$. We want to emphasize that these particular conditions are very natural, since they are satisfied by all sparse pencils in the most relevant families of companion linearizations (in the monomial basis) introduced so far (namely, the Fiedler-like families [1,10,13,20,22,28,40] and the block-Kronecker linearizations [25]).

We follow the developments in [35, p. 624]. We first recall the following notions: Let $\operatorname{trd}(\mathbb{E}/\mathbb{F})$ be the transcendence degree of $\mathbb{E}$ over $\mathbb{F}$, with $\mathbb{F} \subseteq \mathbb{E}$ being a field extension. Since $\{a_0, \ldots, a_k\}$ is a transcendence basis of $\mathbb{F}(a_0, \ldots, a_k)$ over $\mathbb{F}$ [34, p. 317], then $\operatorname{trd}(\mathbb{F}(a_0, \ldots, a_k)/\mathbb{F}) = k + 1$. Instead, given $e_1, \ldots, e_\ell \in \mathbb{E}$ and denoting by $\mathbb{F}(e_1, \ldots, e_\ell)$ the subfield of $\mathbb{E}$ defined by

$$\mathbb{F}(e_1, \ldots, e_\ell) = \left\{ \frac{f(e_1, \ldots, e_\ell)}{g(e_1, \ldots, e_\ell)} \ : \ f, g \in \mathbb{F}[a_0, \ldots, a_\ell], \ g(e_1, \ldots, e_\ell) \neq 0 \right\},$$

then $\operatorname{trd}(\mathbb{F}(e_1, \ldots, e_\ell)/\mathbb{F}) \leq \ell$.

In the proof of Theorems 51 and 53, we use some basic notions of graph theory. Given a pencil $L(\lambda) = [\ell_{st}]$, for $s, t = 1, \ldots, k$, we denote by $\mathcal{D}(L)$ the digraph of $L(\lambda)$. The vertex set of $\mathcal{D}(L)$ is the k-set $V = \{v_1, v_2, \ldots, v_k\}$ and there is an edge (s, t) from the vertex v_s to the vertex v_t if and only if $\ell_{st} \neq 0$, for $1 \leq s, t \leq k$.

Theorem 51 *Let $L(\lambda) = \lambda X + Y$ be a generalized companion pencil for polynomials* (1) *of degree k. Then X and Y altogether have, at least, $3k - 1$ nonzero entries.*

Proof Since $\det L(\lambda) = \alpha q(\lambda)$, with $0 \neq \alpha \in \mathbb{F}$, and the entries of $L(\lambda)$ have degree at most 1 in λ, there must be, at least, k entries of degree 1 in $L(\lambda)$ placed in different rows and columns. Let them be $\lambda x_i + y_i$, with $x_i \neq 0$, for $i = 1, \ldots, k$. By row and column permutation, we can take them to the diagonal positions:

$$\widetilde{L}(\lambda) = \lambda \widetilde{X} + \widetilde{Y} := P_1 \cdot L(\lambda) \cdot P_2 = \begin{bmatrix} \lambda x_1 + y_1 & & * \\ & \ddots & \\ * & & \lambda x_k + y_k \end{bmatrix},$$

with P_1 and P_2 being permutation matrices and $\widetilde{X}, \widetilde{Y} \in \mathbb{F}[a_0, \ldots, a_k]^{k \times k}$. The previous pencil is strictly equivalent to:

$$\widehat{L}(\lambda) = \lambda \widehat{X} + \widehat{Y} := \begin{bmatrix} 1/x_1 & & 0 \\ & \ddots & \\ 0 & & 1/x_k \end{bmatrix} \widetilde{L}(\lambda) = \begin{bmatrix} \lambda + y_1/x_1 & & * \\ & \ddots & \\ * & & \lambda + y_k/x_k \end{bmatrix}.$$

with $\widehat{X}, \widehat{Y} \in \mathbb{F}(a_0, \ldots, a_k)^{k \times k}$. Note that, since $q(\lambda)$ is irreducible over $\mathbb{F}(a_0, \ldots, a_k)$ (Lemma 31), then $\widehat{L}(\lambda)$ is irreducible as well.

Since $\widehat{L}(\lambda)$ is irreducible, by [9, Theorem 3.2.1], $\mathcal{D}(\widehat{L})$ is strongly connected. Furthermore, by [35, Lemma 3], $\mathcal{D}(\widehat{L})$ has a spanning branching and, by [35, Lemma 4], there exists a nonsingular diagonal matrix $D \in M_k(\mathbb{F}(a_0, \ldots, a_k))$ such that

$$\check{L}(\lambda) = \lambda \check{X} + \check{Y} := D \cdot \widehat{L}(\lambda) \cdot D^{-1} \tag{11}$$

has $k - 1$ entries equal to 1 in different positions between $\check{X}$ and $\check{Y}$. Furthermore, from the proof of Lemma 4 in [35], these $k - 1$ entries are not in the main diagonal.

Note that these 1's correspond either to entries in $\check{X}$ or in $\check{Y}$, and that we are not claiming that the full entry in $\check{L}(\lambda)$ is equal to 1. It may be either $\lambda + y$ or $\lambda x + 1$, for some $x, y \in \mathbb{F}(a_0, \ldots, a_k)$.

Note also that the similarity transformation (11) does not affect the diagonal entries of $\widehat{L}(\lambda)$, so $\check{L}(\lambda)$ has

$$\begin{cases} \cdot\ k \text{ diagonal entries of } \check{X} \text{ equal to } 1 \\ \cdot\ k-1 \text{ nondiagonal entries (of either } \check{X} \text{ or } \check{Y}) \text{ equal to } 1. \end{cases} \tag{12}$$

Now, $\det \check{L}(\lambda) = \beta \det L(\lambda) = \beta(\alpha \sum_{i=0}^{k} \lambda^i a_i)$, for some $\beta \in \mathbb{F}(a_0, \ldots, a_k)$. Note also that $\check{L}(\lambda)$ and $L(\lambda)$ have the same number of nonzero entries. Let $e_1, \ldots, e_s$ be the other nonzero entries of $\check{L}(\lambda)$ than the ones in (12) (namely, the remaining nonzero entries in $\check{X}$ and $\check{Y}$, which belong to $\mathbb{F}(a_0, \ldots, a_k)$, counted separately).

Moreover, $a_i \in \mathbb{F}\left(\frac{1}{\beta\cdot\alpha}, e_1, \ldots, e_s\right)$, because $\sum_{i=0}^{k} \lambda^i a_i = \frac{1}{\beta\cdot\alpha} \det \check{L}(\lambda)$, and $\det \check{L}(\lambda)$ is a polynomial in λ whose coefficients belong to $\mathbb{F}(e_1, \ldots, e_s)$. Therefore,

$$\mathbb{F}(a_0, \ldots, a_k) \subseteq \mathbb{F}\left(\frac{1}{\beta\cdot\alpha}, e_1, \ldots, e_s\right). \tag{13}$$

On the other hand, we have

$$\mathbb{F}\left(\frac{1}{\beta\cdot\alpha}, e_1, \ldots, e_s\right) \subseteq \mathbb{F}(a_0, \ldots, a_k). \tag{14}$$

As a consequence of (13) and (14), we get that

$$\mathbb{F}(a_0, \ldots, a_k) = \mathbb{F}\left(\frac{1}{\beta\cdot\alpha}, e_1, \ldots, e_s\right).$$

But

$$k+1 = trd(\mathbb{F}(a_0, \ldots, a_k)/\mathbb{F}) = trd\left(\mathbb{F}\left(\frac{1}{\beta\cdot\alpha}, e_1, \ldots, e_s\right)/\mathbb{F}\right) \le s+1,$$

so $s \ge k$. Therefore,

$$s + 2k - 1 \ge k + 2k - 1 = 3k - 1.$$

In other words, X and Y altogether have, at least, $3k-1$ nonzero entries. □

Some of the nonzero entries of X and Y in Theorem 51 can be placed in the same position. In the Fiedler pencils, for instance, there are only two coincident entries, namely a_{k-1} and a_k, so that any Fiedler pencil has exactly $3k-2$ nonzero entries (see [1,20]). This includes the classical Frobenius companion pencils, $F_1(\lambda)$ and $F_2(\lambda)$, mentioned right before Eq. (2).

Regarding the pencils in the family of quasi-sparse pencils $\mathcal{R}_{1,k}$ introduced in [19], they can have less than $3k-2$ nonzero entries if the coefficients of the polynomial are grouped in pairs of the form $\lambda a_{i+1} + a_i$, as in Proposition 41. This gives a total amount of $2k - 1 + \left\lfloor \frac{k}{2} \right\rfloor$ nonzero entries, which is the smallest number of nonzero entries in the conditions of the statement of Theorems 52 and 53.

Theorem 52 *Let $L(\lambda) = \lambda X + Y$ be a generalized companion pencil for polynomials* (1) *of degree k. If we assume that $L(\lambda)$ has, at least, $2k-2$ nonzero entries with coefficients in $\mathbb{F}$ (namely $\lambda x_{ij} + y_{ij}$ such that $x_{ij}, y_{ij} \in \mathbb{F}$, with at least one of x_{ij} or y_{ij} nonzero), then $L(\lambda)$ has, at least, $2k - 1 + \left\lfloor \frac{k}{2} \right\rfloor$ nonzero entries.*

Proof We essentially follow the proof of Theorem 1 in [35]. The basic idea of the proof is that, to the $2k-2$ nonzero entries of $\lambda X + Y$ with coefficients in $\mathbb{F}$ which are mentioned in the statement, we need to add the $k+1$ coefficients $a_0, \ldots, a_k$, which, in order to minimize the number of nonzero entries, can be grouped in pairs in the same entry (one coefficient in X and the other one in Y). Adding up, we will get the total number of nonzero entries claimed in the statement. To formalize these reasonings, we use some arguments from abstract algebra that involve field extensions. In particular, even though the entries of X and Y belong to the ring $\mathbb{F}[a_0, \ldots, a_k]$, we will consider its field of fractions, $\mathbb{F}(a_0, \ldots, a_k)$, in the proof. Suppose that $L(\lambda)$ has precisely m nonzero entries, counting those of X and Y separately. Let these entries be $e_1, \ldots, e_m$. By the hypothesis in the statement, there are s nonzero entries e_i that are coefficients in $\mathbb{F}$, with $2k-2 \le s \le 4k-4$ (though this bound will not be relevant in the proof). That is, $\{e_1, \dots, e_m\} = \{e_1, e_2, \dots, e_{m-s}\} \cup \{\alpha_1, \dots, \alpha_s\}$, with $\alpha_i \in \mathbb{F}$ (where these sets may include repeated elements).

Note that, by definition of companion pencil, $e_j \in \mathbb{F}[a_0, \dots, a_k] \subseteq \mathbb{F}(a_0, \dots, a_k)$, for $j = 1, \dots, m-s$, and then $\mathbb{F}(e_1, \dots, e_m) = \mathbb{F}(e_1, \dots, e_{m-s}) \subseteq \mathbb{F}(a_0, \dots, a_k)$. On the other hand, since each of the coefficients $a_0, \dots, a_k$ of the polynomial $q(\lambda)$ is the value of a polynomial over $\mathbb{F}$ in the entries of $L(\lambda)$ (that is, $\det L(\lambda) = \alpha q(\lambda)$, with $0 \neq \alpha \in \mathbb{F}$), we have $a_i \in \mathbb{F}(e_1, \dots, e_m) = \mathbb{F}(e_1, \dots, e_{m-s})$, for all $i = 0, \dots, k$, i.e., $\mathbb{F}(a_0, \dots, a_k) \subseteq \mathbb{F}(e_1, \dots, e_{m-s})$.

Therefore $\mathbb{F}(e_1, \dots, e_{m-s}) = \mathbb{F}(a_0, \dots, a_k)$. Finally, from

$$k + 1 = \operatorname{trd}(\mathbb{F}(a_0, \dots, a_k)/\mathbb{F}) = \operatorname{trd}(\mathbb{F}(e_1, \dots, e_{m-s})/\mathbb{F}) \leq m - s$$

we obtain $m \geq k + s + 1$.

Since we are looking for a lower bound on the nonzero entries of $L(\lambda)$, let $m = k+s+1$. Now, in order to minimize the number of nonzero entries in $L(\lambda)$, the nonzero entries $e_1, \dots, e_{m-s}$ can be grouped in pairs (that is, one entry in the matrix X and the other one in the matrix Y, both in the same position) resulting a total amount of $\lceil \frac{m-s}{2} \rceil = \lceil \frac{k+s+1-s}{2} \rceil = \lceil \frac{k+1}{2} \rceil$ nonzero entries in $L(\lambda)$. Besides these ones, there are another $2k-2$ nonzero entries in $\mathbb{F}$, namely the ones mentioned in the statement, which come from $\alpha_1, \dots, \alpha_s$. Adding up, we get, at least, $2k - 2 + \lceil \frac{k+1}{2} \rceil = 2k - 1 + \lfloor \frac{k}{2} \rfloor$ nonzero entries. □

We want to emphasize that all pencils in the Fiedler-like families of companion linearizations for matrix polynomials (namely Fiedler pencils [1,20,22,28], generalized Fiedler pencils [1,13], and generalized Fiedler pencils with repetition [10,40]), as well as the block-Kronecker linearizations in [25], satisfy the conditions of Theorem 52. In particular, the $2k-2$ nonzero entries are of the form: $k-1$ entries equal to 1 and another $k-1$ entries equal to λ (up to scalar constants).

Finally, we have obtained a similar result to Theorem 52 but imposing the restriction that each coefficient a_i, for $i = 0, \dots, k$, of the polynomial (1) appears in just one entry in the generalized companion pencil. In this situation, Proposition 41 provides an explicit description of the nonzero entries, and it is key to prove the result.

Theorem 53 *Let $L(\lambda) = \lambda X + Y$ be a generalized companion pencil for polynomials* (1) *of degree k. If we assume that each coefficient a_i, for $i = 0, \dots, k$, appears in just one entry of $L(\lambda)$, then $L(\lambda)$ has, at least, $2k - 1 + \lfloor \frac{k}{2} \rfloor$ nonzero entries.*

Proof Throughout the proof, a set of k nonzero entries in $L(\lambda)$ located in different rows and columns is referred to as a *composite k-cycle* (of the digraph $\mathcal{D}(L)$), following the nomenclature in [26, p. 258]. The label of the edge is the nonzero entry of $L(\lambda)$ corresponding to this edge. We will say that an edge "contains" a coefficient a_i if the label contains the coefficient a_i. For brevity, we will also say that a cycle "contains" a coefficient a_i if the entry containing this coefficient is an edge of the cycle.

We are going to prove first that, in the conditions of the statement, for a given $0 \leq j \leq k$, there is, at least, one composite k-cycle with an edge containing a_j, and such that the remaining edges do not contain any other coefficient a_i. For this, let $z_j a_j + w_j$ be the entry containing a_j, where z_j and w_j are polynomials of degree at most 1 in λ, like in Proposition 41. Then, spanning across the row containing $z_j a_j + w_j$ we get

$$\det L(\lambda) = \pm(z_j a_j + w_j) \det L_j(\lambda) + q_j(\lambda),$$

where $q_j(\lambda)$ is a polynomial not containing a_j, and $L_j(\lambda)$ is the subpencil of $L(\lambda)$ obtained after removing the row and column containing $z_j a_j + w_j$. Now, let $0 \leq i \leq k$ with $i \neq j$, and let us consider the entry $z_i a_i + w_i$ containing a_i, where z_i, w_i are, again, polynomials of degree at most 1 in λ, as in Proposition 41. If the edge corresponding to this entry is part of a composite k-cycle including the edge of $z_j a_j + w_j$, then $z_i a_i + w_i$ is in $L_j(\lambda)$, and spanning the determinant of this pencil across the row containing this entry, we get

$$\det L(\lambda) = \pm(z_j a_j + w_j)\,(\pm(z_i a_i + w_i) \det L_{ij}(\lambda) + q_{ij}(\lambda)) + q_j(\lambda), \tag{15}$$

where $q_{ij}(\lambda)$ is a polynomial in λ not containing a_i, and $L_{ij}(\lambda)$ is the subpencil of $L_j(\lambda)$ obtained after removing the row and column containing $z_i a_i + w_i$. Since the product $a_i a_j$ only appears in the first summand of the product $(z_j a_j + w_j)(z_i a_i + w_i) \det L_{ij}(\lambda)$ in (15), and $L(\lambda)$ is a generalized companion pencil, it must be $\det L_{ij}(\lambda) \equiv 0$. As a consequence, all cycles containing the edges corresponding to

$z_j a_j + w_j$ and $z_i a_i + w_i$ cancel out. This means that the term $\lambda^j a_j$ in $\det L(\lambda)$ must come from, at least, another composite k-cycle not containing any other coefficient a_i, for $i \neq j$, in a different edge.

Now, among the composite k-cycles containing a_k and not containing any other coefficients a_i in other edges, we choose one and denote it by $\mathcal{C}_k$. We are going to see that, if all composite k-cycles containing a_0 and not containing a_k have a common edge with $\mathcal{C}_k$, then $L(\lambda)$ would not be a generalized companion pencil. For this, we assume, by contradiction, that all composite k-cycles containing a_0 and not containing a_k contain the common edge labelled as $\lambda\beta + \alpha$ in $\mathcal{C}_k$. Let $\alpha_0 a_0 + w_0$ and $\lambda\beta_k a_k + w_k$, with $\alpha_0, \beta_k \in \mathbb{F} \setminus \{0\}$, be the entries containing a_0 and a_k, respectively. Note that, by Proposition 41, a_k can only appear multiplied by λ (and a_0 must appear without the factor λ instead). Then

$$\det L(\lambda) = (\lambda\beta + \alpha)\left((\alpha_0 a_0 + w_0)q_0(\lambda) + (\lambda\beta_k a_k + w_k)q_k(\lambda)\right) + r(\lambda),$$

where $q_0(\lambda), q_k(\lambda)$, and $r(\lambda)$ are polynomials in λ, with $r(\lambda)$ containing neither a_0 nor a_k, $q_k(\lambda)$ does not contain a_0, and has degree at most $k-2$. As a consequence, it must be $\beta \neq 0$, since, otherwise, the term containing a_k in $\det L(\lambda)$ would not have degree k. But, then, $\det L(\lambda)$ would contain a term of the form $\lambda\beta a_0$, so $L(\lambda)$ would not be a generalized companion pencil.

Now, let $\mathcal{C}_0$ be a composite k-cycle containing a_0 and no other a_i in a different edge (entry), and let us assume that this cycle contains d common edges with $\mathcal{C}_k$. Then, there must be another d additional edges to replace these ones, and none of these edges (entries) contain any other a_i, except, maybe, the entry containing a_0. In order to see this, let us assume that the common d edges correspond to the first diagonal positions of $L(\lambda)$ (otherwise, we can take them to these positions by row and column permutation). These edges correspond to d loops, $\ell_1, \dots, \ell_d$, in vertices $v_1, \dots, v_d$ of $\mathcal{D}(L)$. Since none of these loops can be in all k-cycles containing a_0 and not containing a_k (since, as we have seen in the previous paragraph, otherwise $L(\lambda)$ would not be a generalized companion pencil), there must be a composite k-cycle, $\mathcal{C}_0^{(i)}$, containing a_0 and not containing ℓ_i, for each $1 \leq i \leq d$ (and not containing either any other a_j in a different edge). Therefore, for each vertex v_i, with $1 \leq i \leq d$, there must be another two edges, one inner edge, and one outer edge. The minimum number of edges that we need to add in order to fulfill these requirements can be obtained with a d-cycle involving all vertices $v_1, \dots, v_d$ (that is, a cycle of the form $(v_1, v_2), (v_2, v_3), \dots, (v_{d-1}, v_d), (v_d, v_1)$), which gives a total amount of d new edges. Therefore, we need, at least, d additional edges to replace these d loops. Adding up to the remaining $k-d$ edges of $\mathcal{C}_0$, there must be another k entries, other than those in $\mathcal{C}_k$, corresponding to composite k-cycles containing a_0 and not containing a_k.

So far, we have proved that $L(\lambda)$ contains, at least, $2k$ nonzero entries. One of these entries contains a_0 and another one contains a_k, but the remaining ones do not contain any other coefficient a_i. These other coefficients, however, must be in $L(\lambda)$. The smallest number of additional entries that we can add to these $2k$ entries in order to incorporate these coefficients is $\lceil \frac{k-3}{2} \rceil$. This can be done by adding a_1 to the entry containing a_0 and a_{k-1} to the entry containing a_k, and grouping in pairs the remaining $k-3$ coefficients, as in (7)–(8). The identity $\lceil \frac{k-3}{2} \rceil = \lfloor \frac{k}{2} \rfloor - 1$ concludes the proof. □

In the following example, we show a sparse companion pencil, $L_1(\lambda)$, in the family $\mathcal{R}_{1,k}$ introduced in [19], and another companion pencil, $L_2(\lambda)$, where each coefficient of the polynomial appears only once, which is not sparse (and not in $\mathcal{R}_{1,k}$ either). This pencil $L_2(\lambda)$ has the additional property that the only composite k-cycle of degree k containing a_k has a common edge with a composite k-cycle containing a_0.

Example 51 *Let $q(\lambda) = \sum_{i=0}^{4} \lambda^i a_i$ be a polynomial of degree 4. Let us consider the following matrix pencils $L_1(\lambda)$ and $L_2(\lambda)$:*

$$L_1(\lambda) = \begin{bmatrix} \lambda & -1 & 0 & 0 \\ 0 & \lambda & -1 & 0 \\ 0 & 0 & \lambda & -1 \\ \lambda a_1 + a_0 & \lambda a_2 & 0 & \lambda a_4 + a_3 \end{bmatrix} \quad \text{and} \quad L_2(\lambda) = \begin{bmatrix} \lambda & \lambda a_1 + a_0 & 1 & -1 \\ 0 & \lambda & 0 & 1 \\ \lambda & \lambda & \lambda + 1 & \lambda \\ 1 & -\lambda + 2 - a_2 & 1 & \lambda a_4 + a_3 \end{bmatrix}.$$

The coefficients $a_0, \dots, a_4$ appear only once in both $L_1(\lambda)$ and $L_2(\lambda)$. Moreover, both pencils are generalized companion because $\det L_1(\lambda) = \det L_2(\lambda) = \sum_{i=0}^{4} \lambda^i a_i$. However, only the pencil $L_1(\lambda)$ is sparse, having $2 \cdot 4 - 1 + \lfloor \frac{4}{2} \rfloor = 9$ nonzero entries. Note that, in $L_1(\lambda)$, there is only one composite 4-cycle containing each coefficient a_i, for $1 \leq i \leq 4$. In $L_2(\lambda)$, the only composite 4-cycle of degree 4 containing

a_4 is $(1,1),(2,2),(3,3),(4,4)$, which has the edge $(3,3)$ in common with the following composite 4-cycle containing a_0: $(1,2)-(2,4)-(4,1),(3,3)$. However, there are other composite 4-cycles containing either a_0 or a_4.

6 Conclusions and open questions

In this note, we have introduced the notion of *generalized companion pencil*, $L(\lambda) = \lambda X + Y$, for scalar polynomials as in (1), which extends the notion of companion matrix A of monic scalar polynomials (with $X = I$ and $Y = -A$), and the notion of *companion pencil* by allowing the coefficients X and Y to contain entries in the ring of polynomials in the coefficients $a_0, \ldots, a_k$. We have proved that all generalized companion pencils of a given polynomial $q(\lambda)$ as in (1) have the same Smith canonical form over $\mathbb{F}(a_0, \ldots, a_k)$, namely $\mathrm{diag}(I_{k-1}, \frac{1}{a_k} q(\lambda))$, and that they are all nonderogatory.

We have seen that, if we impose the condition that each coefficient a_i of the polynomial (1) appears only once in $L(\lambda)$, then the entries of $L(\lambda)$ have a very specific form, and we have also seen that, under this condition, the smallest number of nonzero entries of $L(\lambda)$ is $2k - 1 + \lfloor \frac{k}{2} \rfloor$ (Theorem 53). This is also the smallest number of nonzero entries of $L(\lambda)$ if it contains $k-1$ entries equal to 1 and another $k-1$ entries equal to λ (up to nonzero constant factors, see Theorem 52), as it happens in most of the families of companion pencils introduced so far in the literature. For generalized companion pencils without restrictions, we have proved that the number of nonzero entries in X and Y, altogether, is at least $3k-1$. However, some of these nonzero entries may be in the same position in $L(\lambda)$.

As an open question, it remains to determine the smallest possible number of nonzero entries in an arbitrary generalized companion pencil for polynomials of degree k. We conjecture that this number is $2k - 1 + \lfloor \frac{k}{2} \rfloor$.

Acknowledgements. We are very much indebted to an anonymous referee for a very careful reading of the manuscript and for many helpful suggestions that allowed us to improve significantly the original version. We also thank a second referee for suggesting the inclusion of several references on companion matrices for polynomials expressed in other bases than the monomial basis, like [5,8,33,36,37].

References

1. E. N. Antoniou, S. Vologiannidis. A new family of companion forms for polynomial matrices. Electron. J. Linear Algebra, 11 (2004) 78–87.
2. J. Aurentz, T. Mach, R. Vandebril, D. S. Watkins. Fast and backward stable computation of roots of polynomials. SIAM J. Matrix Anal. Appl. 36 (2015) 942–973.
3. J. Aurentz, T. Mach, L. Robol, R. Vandebril, D. S. Watkins. Fast and backward stable computation of roots of polynomials, part II: backward error analysis; companion matrix and companion pencil. SIAM J. Matrix Anal. Appl. 39 (2018) 1245–1269.
4. J. Aurentz, T. Mach, L. Robol, R. Vandebril, D. S. Watkins. Fast and backward stable computation of the eigenvalues of matrix polynomials. Math. Comp. 88 (2019) 313–347.
5. S. Barnett. A companion matrix analogue for orthogonal polynomials. Linear Algebra Appl. 12 (1975) 197–208.
6. S. Barnett. Congenial matrices. Linear Algebra Appl. 41 (1981) 277–298.
7. D. Bini, L. Gemignani, V. Pan. Fast and stable QR eigenvalue algorithms for generalized companion matrices and secular equations. Numer. Math., 100 (2005) 373–408.
8. J. P. Boyd. Finding the zeros of a univariate equation: Proxy rootfinders, Chebyshev interpolation, and the companion matrix. SIAM Rev. 55 (2013) 375–396.
9. R. Brualdi, H. Ryser. Combinatorial Matrix Theory. Cambridge University Press, Cambridge, 1991.
10. M. I. Bueno, K. Curlett, S. Furtado. Structured strong linearizations from Fiedler pencils with repetition I. Linear Algebra Appl. 460 (2014) 51-80.
11. M. I. Bueno, K. Curlett, S. Furtado. Structured linearizations from Fiedler pencils with repetition II. Linear Algebra Appl. 463 (2014) 282–321.
12. M. I. Bueno, F. De Terán. Eigenvectors and minimal bases for some families of Fiedler-like linearizations. Lin. Multilin. Algebra, 62 (2014) 39–62.
13. M. I. Bueno, F. De Terán, F. M. Dopico. Recovery of eigenvectors and minimal bases of matrix polynomials from generalized Fiedler linearizations. SIAM J. Matrix Anal. Appl. 32 (2011) 463–483.
14. M. I Bueno, F. M. Dopico, J. Pérez, R. Saavedra, B. Zykovsi. A unified approach to Fiedler-like pencils via strong block minimal bases pencils. Linear Algebra Appl. 547 (2018) 45–104.
15. C. Carstensen. Linear construction of companion matrices. Linear Algebra Appl. 149 (1991) 191–214.

16. R. M. Corless. Generalized companion matrices in the Lagrange basis. In: L. Gonzalez-Vega and T. Recio (eds.). Proceedings EACA, 317–322, June 2004.
17. R. M. Corless. On a generalized companion matrix pencil for matrix polynomials expressed in the Lagrange basis. In: D. Wang, L. Zhi (eds). Symbolic-Numeric Computation. Trends in Mathematics. Birkhuser, Basel, 2007.
18. D. Day, L. Romero. Roots of polynomials expressed in terms of orthogonal polynomials. SIAM J. Numer. Anal. 43 (2006) 1969–1987.
19. F. De Terán, C. Hernando. A class of quasi-sparse companion pencils. In: D. A. Bini et al. (eds.). Structured Matrices in Numerical Linear Algebra. INdAM series, 157–179, Springer, 2019.
20. F. De Terán, F. M. Dopico, D. S. Mackey. Fiedler companion linearizations and the recovery of minimal indices. SIAM J. Matrix Anal. Appl. 31 (2010) 2181–2204.
21. F. De Terán, F. M. Dopico, D. S. Mackey. Palindromic companion forms for matrix polynomials of odd degree. J. Comput. Appl. Math. 236 (2011) 1464–1480.
22. F. De Terán, F. M. Dopico, D. S. Mackey. Fiedler companion linearizations for rectangular matrix polynomials. Linear Algebra Appl. 437 (2012) 957–991.
23. F. De Terán, F. M. Dopico, D. S. Mackey. Spectral Equivalence of matrix polynomials and the Index Sum Theorem. Linear Algebra Appl. 459 (2014) 264-333.
24. L. Deaett, J. Fischer, C. Garnett, K. N. Vander Meulen. Non-sparse companion matrices. Electron. J. Linear Algebra, 35 (2019) 223–247.
25. F. M. Dopico, P. Lawrence, J. Pérez, P. Van Dooren. Block Kronecker linearizations of matrix polynomials and their backward errors. Numer. Math. 140 (2018) 373–426.
26. B. Eastman, I.-J. Kim, B. L. Shader, K.N. Vander Meulen. Companion matrix patterns. Linear Algebra Appl. 436 (2014) 255-272.
27. B. Eastman, K. N. Vander Meulen. Pentadiagonal companion matrices. Spec. Matrices, 4 (2016) 13–30.
28. M. Fiedler. A note on companion matrices. Linear Algebra Appl. 372 (2003) 325-331.
29. S. Fortune. An iterated eigenvalue algorithm for approximating roots of univariate polynomials J. Symbolic Computation (2002) 33, 627–646.
30. C. Garnett, B. L. Shader, C. L. Shader, P. van den Driessche. Characterization of a family of generalized companion matrices. Linear Algebra Appl. 498 (2016) 360-365.
31. L. Gemignani. Structured matrix methods for polynomial root-finding. In ISSAC07, 175–180, 2007, ACM.
32. I. Gohberg, P. Lancaster, L. Rodman. Matrix Polynomials. Academic Press, New York, 1982.
33. I. J. Good. The colleague matrix, a Chebyshev analogue of the companion matrix. Quart. J. Math Oxford Ser. (2) 12 (1961) 61–68.
34. T. W. Hungerford. Algebra. GTM 73, Springer, 1980.
35. C. Ma, X. Zhan. Extremal sparsity of the companion matrix of a polynomial. Linear Algebra Appl. 438 (2013) 621-625.
36. J. Maroulas, S. Barnett. Polynomials with respect to a general basis I. Theory. J. Math. Anal. Appl. 72 (1979) 177-194.
37. J. Nakatsukasa, V. Noferini. On the stability of computing polynomial roots via confederate linearizations. Math. Comp. 85 (2016) 2391–2425.
38. U. Prells, M. I. Friswell, S. D. Garvey. Use of geometric algebra: compound matrices and the determinant of the sum of two matrices. Proc. R. Soc. Lond. A 453 (2003) 273–285.
39. P. Van Dooren, P. Dewilde. The eigenstructure of an arbitrary polynomial matrix: computational aspects. Linear Algebra Appl. 50 (1983) 545–579.
40. S. Vologiannidis, E. N. Antoniou. A permuted factors approach for the linearization of polynomial matrices. Math. Control Signals Syst. 22 (2011) 317–342.
41. W. Werner. A generalized companion matrix of a polynomial and some applications. Linear Algebra Appl. 55 (1983) 19–36.